

Assessing the direction of climate interactions by means of complex networks and information theoretic tools

J. I. Deza,^{1,*} M. Barreiro,² and C. Masoller¹

¹*Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya,
Colom 11, E-08222, Terrassa, Barcelona, Spain.*

²*Instituto de Física, Facultad de Ciencias, Universidad de la República, Iguá 4225, Montevideo, Uruguay.*
(Dated: February 5, 2015)

An estimate of the net direction of climate interactions in different geographical regions is made by constructing a directed climate network from a regular latitude-longitude grid of nodes, using a directionality index (DI) based on conditional mutual information. Two datasets of surface air temperature anomalies—one monthly-averaged and another daily-averaged—are analyzed and compared. The network links are interpreted in terms of known atmospheric tropical and extra-tropical variability patterns. Specific and relevant geographical regions are selected, the net direction of propagation of the atmospheric patterns is analyzed and the direction of the inferred links is validated by recovering some well-known climate variability structures. These patterns are found to be acting at various time-scales, such as atmospheric waves in the extratropics or longer range events in the tropics. This analysis demonstrates the capability of the DI measure to infer the net direction of climate interactions and may contribute to improve the present understanding of climate phenomena and climate predictability. The work presented here also stands out as an application of advanced tools to the analysis of empirical, real-world data.

PACS numbers: 92.70.Gt, 89.70.-a, 87.18.Sn

Keywords: Climate Networks, Directionality, Information Theory

Information-theoretic tools are used to construct *directed* climate networks from time-series analysis of observed climatological data. Specifically, surface air temperature (SAT) anomalies are considered. Two datasets—one monthly-averaged and another daily-averaged—are used. Directed links among the network nodes are defined via an analysis of the net direction of information transfer. A predictability measure—based on conditional mutual information—quantifying the amount of information in a time-series $x(t)$, contained in τ time units in the past of another time series, $y(t)$ is used. The resulting directed network is then studied and a full agreement with state-of-the-art knowledge in climate phenomena has been found, validating this methodology for inferring the net directionality of climate interactions, directly from the data. No weather assumptions or models are made, except for the appropriate setting of the parameter τ which is sensible to the shorter or longer autocorrelation of the time series.

I. INTRODUCTION

Network theory is a well-known framework for describing complex systems composed of many interacting components [1–4]. Many systems can be straightforwardly represented in terms of a well-defined set of nodes coupled among them via links that have a clear physical interpretation. This is the case, for example, of airport networks [5], of social interactions [6] or the Internet [7], just to name some. In other systems it is not clear how to define the relevant nodes, and/or there is no obvious interaction that can be used to define links. An example of this situation is the Earth climate system, in which the lattice of grid points from measurements or models, is defined to be the network set of nodes, and it depends on the resolution of the dataset analyzed. Many climatological fields—such as surface air temperature (SAT) or the geopotential height (GH) at a certain pressure level—can be used to define links via an analysis of significant correlations [8–10].

Climate networks (CNs), which represent the statistical similarity structure of spatio-temporal resolved climatological variables, depend on the definition of nodes and links [11]. Their regular spatial sampling results in a small-world topology [12], and thus, a careful interpretation of the inferred network is required. Nevertheless, CNs have been successfully employed to analyze climate features including the global connectivity [13–15], the identification of community structures [16, 17] and the study of the possible collapse of the meridional overturning circulation (MOC) [18] on the north Atlantic, among others.

A relevant drawback of this correlation analysis,

* juan.ignacio.deza@upc.edu.

which uses symmetric similarity measures (such as cross-correlation or mutual information), is that it yields non-directed networks where the presence of a link reveals inter-dependency but the direction associated (if any) of the underlying interaction is not established. For improving the understanding of climate phenomena and its predictability, it is of foremost importance not only to be able to infer the presence of a link between two nodes, but also, to infer the direction of this interaction.

A path to overcome this limitation is by constructing weighted climate networks, where the weight of each link is composed of two numbers: the correlation strength between the two nodes and the time delay which maximizes this correlation strength. The sign of the time delay gives information about the direction of the link. Using this approach, the high sensitivity of the network links to El-Niño events was demonstrated, even in geographical regions far from the Pacific ocean [19, 20].

An alternative approach for assessing the directionality of climate interactions involves the use of Granger Causality. In [21] it was argued that the inclusion of several variability patterns like NAO (North Atlantic Oscillation), PDO (Pacific Decadal Oscillation), ENSO (El Niño–Southern Oscillation), and NPI (North Pacific Index)—which occur naturally and explain an important part of the global atmospheric variability—into a nonlinear network-like prediction method, largely improves the predictability of global temperature over seasonal time scales. This is so even when the time scales of the patterns used of in the order of decades. The study suggested a causal directional influence among these major oceanic and atmospheric modes and global temperature variability, not only over their own time-scales but also over much shorter ones. Granger causality has also been used to test interdependence among ENSO and the Indian monsoon [22]. A non-symmetric bidirectional and even alternating character of coupling was found that extends previous knowledge about the presence of negative correlation and intervals of phase synchronicity between the processes.

A third approach for directionality detection is based on information-theoretic measures [23]. For example, in [24] information transfer from larger to smaller time scales was detected in daily-averaged SAT time series as the causal influence of the phase of slow oscillatory phenomena with periods of about 6-11 years in the amplitudes of the variability, characterized by smaller timescales, from a few months to 4-5 years.

The reliability of directed climate networks detected by bivariate nonlinear methods based on information theory, compared to those generated by linear Granger causality analysis, was studied in [25]. Several algorithms for estimating transfer entropy with a wide range of parameter choices were considered. As transfer entropy is a special case of the conditional mutual information, it reduces to Granger causality for linear Gaussian processes, and it usually requires longer time-series for accurate estimation. It was shown that all the causality methods consid-

ered provided reproducible estimates of climate directed interactions, with the linear method outperforming the nonlinear ones in terms of reliability.

In the present paper, directed climate networks are constructed by using a predictability measure which, to the best of our knowledge, has not yet been applied to climatological data. This bivariate analysis quantifies the amount of information contained in a time-series, $x(t)$, about the past of another time series, $y(t)$, τ time units before. More specifically, the *directionality index* (DI) is used, which quantifies the net direction of information flow between two time series, and has been successfully employed for the analysis of numerical data generated from coupled oscillators, and empirical data from cardiorespiratory recordings and electroencephalographic recordings, to name just a few [26–28].

The objective of this work is to demonstrate that the DI methodology for inferring the net direction of information transfer indeed unveils climatologically relevant phenomena. Clear patterns of variability are uncovered, both in the tropics and in the extra-tropical regions, and their direction of propagation is shown to be in good agreement with the current understanding of main climate phenomena. The method is shown to work with both monthly- and daily-averaged data, and furthermore, well-defined atmospheric patterns are uncovered on daily-averaged data, which are not seen with monthly-averaged data.

The paper is organized as follows. Section II presents the data analyzed and the method used for constructing directed climate networks. The statistical significance test is also discussed. Section III presents the results and provide a comparison between directed and non-directed climate networks. Relevant patterns of global atmospheric variability are interpreted from the maps obtained through the DI analysis. Finally Sec. IV presents the conclusions.

II. DATA AND METHODS

A. Data

We consider two datasets, both corresponding to SAT anomalies from the reanalysis of the National Center for Environmental Prediction/National Center for Atmospheric Research, (NCEP/NCAR)[29]. The data covers a regular grid over the Earth’s surface with latitudinal and longitudinal resolution of 2.5° , resulting in $N = 10226$ grid points or network nodes.

The first dataset corresponds to monthly-averaged SAT data. Since the data cover the period from January 1949 to December 2013, in each node there is a time series of 780 data points. The SAT anomalies are calculated as the actual temperature values minus the monthly average, and they are normalized by the standard deviation. Each time-series is linearly detrended.

The second dataset consists of daily-averaged SAT

from the same source, with the same spatial resolution, and covering the same period of time (the length of each time series is 23725 data points). In this case, for calculating the anomalies, the daily average has been subtracted from the actual temperature value, and the leap days have been discarded. The time series are also detrended and normalized by the standard deviation.

B. Directionality measure

We consider the SAT anomalies time series in two nodes, $X(t)$ and $Y(t)$, which are characterized by probability distribution functions (PDFs) p_X , p_Y , and by their joint PDF, p_{XY} . To assess the directionality of the link between these two nodes the *directionality index* (DI) as defined in [26, 30] is used:

$$DI_{XY}(\tau) = \frac{I_{XY}(\tau) - I_{YX}(\tau)}{I_{XY}(\tau) + I_{YX}(\tau)}, \quad (1)$$

where $I_{XY}(\tau)$, $I_{YX}(\tau)$ (conditional mutual information) are defined as:

$$\begin{aligned} I_{XY}(\tau) &= I(X; Y|X_\tau) \\ &= H(X|X_\tau) + H(Y|X_\tau) - H(X, Y|X_\tau); \quad (2) \\ I_{YX}(\tau) &= I(Y; X|Y_\tau) \\ &= H(Y|Y_\tau) + H(X|Y_\tau) - H(Y, X|Y_\tau), \quad (3) \end{aligned}$$

with $X_\tau = X(t - \tau)$, $Y_\tau = Y(t - \tau)$ and $H(X|Y)$ being the conditional entropy [26, 30].

In terms of three time series $X(t)$, $Y(t)$ and $Z(t)$, $I(X; Y|Z)$ indicates the amount of information shared between $X(t)$ and $Y(t)$, given the effect of $Z(t)$ over $Y(t)$. In this way, to assess the information transfer from $X(t)$ to $Y(t)$, $Z(t)$ is replaced by the past of time-series $X(t)$. Thus, $I(X; Y|X_\tau)$ quantifies the amount of information shared between $X(t)$ and $Y(t)$, given the influence of $X(t - \tau)$ over $Y(t)$. Analogously, to assess the information transfer from Y to X , $Z(t)$ is replaced by the past of $Y(t)$. This special case of conditional mutual information is also referred as *transfer entropy*.

The directionality index, DI_{XY} , then quantifies the *net* information flow. From the definition of DI_{XY} , Eq.(1), it is clear that $DI_{XY} = -DI_{YX}$. Also, $-1 \leq DI_{XY} \leq 1$: $DI_{XY} = 1$ if and only if $I_{XY} \neq 0$, $I_{YX} = 0$ (*i.e.*, the information flow is $X \rightarrow Y$ and there is no back coupling $Y \rightarrow X$) and $DI_{XY} = -1$ if and only if $I_{XY} = 0$, $I_{YX} \neq 0$ (*i.e.*, the information flow is $Y \rightarrow X$ and there is no back coupling $X \rightarrow Y$).

Naturally, $\tau > 0$ is a parameter that has to be tuned appropriately to the time-scales relevant to the particular dataset. If τ is too small $DI_{XY}(\tau)$ will capture short time scale directionality, and may fail if the time series behave too similarly on those time scales as they do if they are subjected to the same external forcing. On the other hand, if τ is too large, larger than the decorrelation time of the time-series, the effect of the past X over Y

(and of Y over X) will be negligible and $DI_{XY}(\tau)$ will be a small and in principle random value. In the next section the criterion employed for accessing the statistical significance of the DI values will be discussed, and in section III, a thorough study of the effect of varying the parameter τ will be performed. It is worth noticing that in the two datasets considered, $\tau = 1$ has a different meaning: in the monthly-averaged SAT time-series, the minimum value of τ is one month, while in the daily-averaged data, the minimum value of τ is one day.

To compute the PDFs associated to each time series, and the joint PDFs a 10-bin histograms of values is used. An alternative approach is to use a symbolic transformation known as *ordinal analysis* [26, 31–33]. A preliminary study suggests that using ordinal analysis offers as an additional advantage the possibility of finding the directionality of the links at different time-scales; this study is still in progress.

It is important to note a drawback of using the directionality index for network construction: it does not distinguish indirect from direct information transfer. Given three time-series, X , Y and Z , positive and significant DI values $I_{XY} > 0$ and $I_{XZ} > 0$ might also give a significant value of I_{YZ} —either positive or negative—even if there is no “direct” information in Y about the future or the past of Z , as the information will have been “indirectly” contained in X .

III. RESULTS

Firstly, the results of the statistical significance analysis of link directionality, when using monthly-averaged SAT datasets, are presented. Monthly datasets are used in order to compare with the undirected networks reported in [33]. Afterwards, the results of the analysis of the daily-averaged datasets are presented. They are consistent with the monthly datasets results, furthermore uncovering additional patterns of atmospheric variability, not observed in monthly data as the averaging procedure filters high frequency and variability.

A. Statistical significance analysis

To address the significance of DI values, 100 surrogates were generated using the bootstrap (BS) algorithm [34]. The null hypothesis considered for MI is that the processes are independent from each other. The null hypothesis for DI is that there is no preferential direction of the interaction.

The BS algorithm randomly resamples with replacement from the original datasets using blocks of data of approximately the size of the autocorrelation time of the time series, and then computes the estimators (MI and DI) from the resampled data. Doing so, both the statistics (histogram) and the power spectrum of the original time series are approximately preserved. In this way,

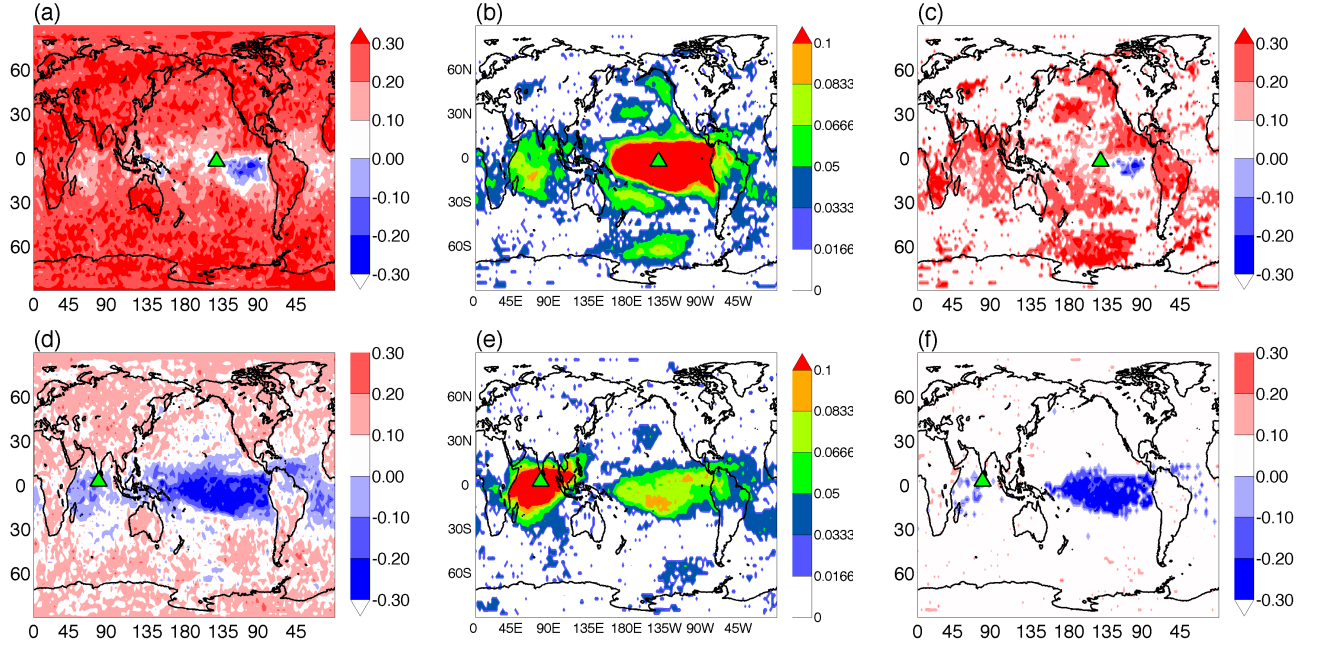


FIG. 1. (Color Online). Procedure of constructing significant directionality maps from raw DI calculations. In panels (a) and (d) the unfiltered DI maps are shown for two nodes, one in the central Pacific ocean and one in the Indian ocean, indicated with triangles. In panels (b) and (e) only the statistically significant MI values are shown. These results are combined in (c) and (f) where only the links that have both, statistically significant MI and DI values are shown.

two different empirical distributions (one-tailed for MI and two-tailed for DI), were obtained for each link, and significance thresholds were extracted from them.

Afterwards, for each link, the MI value was calculated from the original datasets and a significance test was applied. A MI value is considered significant if: $MI > \Theta_{MI}$, where Θ_{MI} is the threshold derived from the bootstrap MI distribution. In addition—for the MI significant links only—the DI value computed from the original dataset was compared to the DI thresholds Θ_{DI}^+ and Θ_{DI}^- . The DI value was considered significant if $DI > \Theta_{DI}^+$ or if $DI < \Theta_{DI}^-$. This two-step significance test was performed with the scope of assessing firstly the presence of a link, and afterwards its directionality.

A graphical explanation of the full procedure is shown in Fig. 1 where, for two nodes (one over the Pacific and one over the Indian ocean) the unfiltered DI maps are displayed in panels (a,d), the significant MI values are displayed in panels (b,e) and finally the DI significant values for the MI significant links are displayed in panels (c,f). The DI maps show positive DI values in red, which mean outgoing links, while the incoming links are shown in blue.

The significance thresholds, Θ_{MI} , Θ_{DI}^+ and Θ_{DI}^- , extracted from the BS surrogates, were computed as $\Theta_{MI} = \mu_{MI} + 3 \times \sigma_{MI}$, $\Theta_{DI}^+ = \mu_{DI} + 3 \times \sigma_{DI}$, and $\Theta_{DI}^- = \mu_{DI} - 3 \times \sigma_{DI}$, with μ and σ being the mean and standard deviation of the corresponding MI and DI bootstrap distributions. The process is illustrated in Fig.

2, panels (a) and (b). It can be noticed that these distributions are bell-shaped, which justifies the use of classical statistics. Also notice from panel 2(b) that for a sufficiently large number of BS surrogates, the distribution tends to be centered around a rather small DI value, which motivates the null hypothesis of no preferential direction of the interaction.

The more time and memory consuming 99nth percentile test was also used, and yielded similar results.

In order to visualize the results of the significance criterion employed, in Fig. 2 panels (c) and (d) display—as an example—the DI *vs* MI values of all the links of the two nodes, one located in the central Pacific and the other in the Indian ocean (the nodes are the same as in Fig. 1). The black dots indicate the disregarded links (either because MI is not significant, or because MI is significant but DI is not). Significant links are plotted in red (outgoing links, $DI > \Theta_{DI}^+$) and in blue (incoming links, $DI < \Theta_{DI}^-$).

In Fig. 2, panels (c) and (d) also show that high MI values do not imply high DI values. In Fig. 2 (d) one can notice that most of the blue dots are located in a narrow range of MI values while they are distributed in terms of DI values. An inspection of panel (f) in Fig. 1 shows that the blue links come to the node in the Indian Ocean from a well-defined region in the central Pacific Ocean. On the other hand, one can notice in Fig. 2 (d) that the few red dots are more distributed in the MI, DI plane, and in Fig. 1 (f) the red outgoing links connect the node

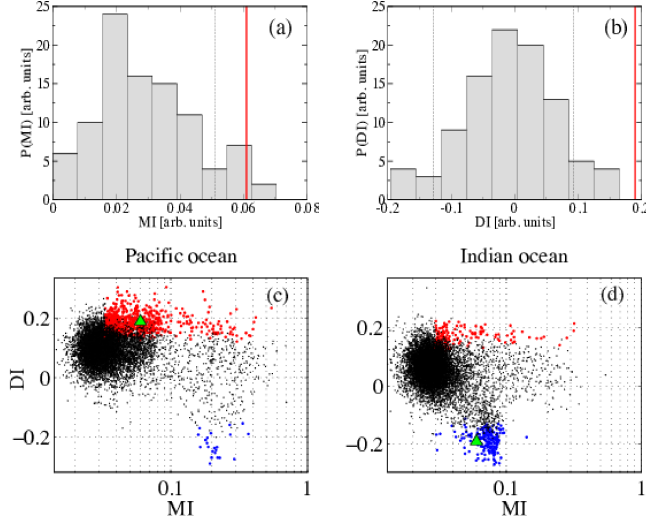


FIG. 2. (Color Online). (a), (b) Histograms of the MI and DI values ($\tau = 1$ month) computed from 100 BS surrogates, the dashed vertical lines indicate the significance thresholds and the solid lines (red online) indicate the MI and DI values computed from the original data. (c), (d) Plot of DI ($\tau = 1$ month) vs. MI for all the links of the two nodes in the Pacific and Indian oceans considered in Fig. 1. Incoming links are indicated in blue and outgoing links, in red. Significant links are plotted in red or blue dots; the black dots indicate the disregarded links (either because MI is not statistically significant, or because MI is significant but DI is not). The triangles indicate the (MI, DI) values of the particular link between the two nodes.

in the Indian Ocean to various regions on Earth.

B. Analysis of monthly-averaged SAT anomalies

1. Influence of the parameter τ

As stated in the methodology the correct choice of the value of τ is necessary for obtaining consistent results. As τ can be only an integer, using monthly averaged data, its minimum value will be of one month. In the tropical areas the influence of the ocean on the surface air temperature is a dominant characteristic. Moreover, because of the large heat capacity of water and the ocean's dynamics, the sea surface temperature (SST) anomalies vary in the scale of months. Calculating DI for a point in the central pacific (NINO3 area) for different values of τ yields the results shown in Fig. 3. The point considered is the same as in Fig. 1 (a-c) and moreover, the panel 1 (c) is the same as panel 3 (a).

For $\tau = 1$ Fig.3 (a) shows the central Pacific influenced by (in blue) the eastern Pacific and influencing (in red), presumably through atmospheric teleconnections, the global tropics and the extratropical Pacific ocean. However, as τ grows the number of significant connec-

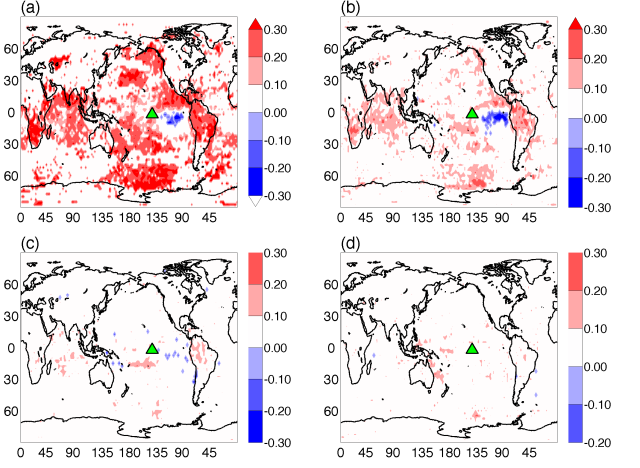


FIG. 3. (Color Online). Effect of τ on tropical areas using monthly-averaged data. In this case a point in central pacific (the same point used in the top row of Fig. 1 and in [33]) is considered. The values of τ are: (a) 1 month, (b) 3 months, (c) 6 months, (d) 12 months. Notice the decorrelation of the time series for large τ . Incoming links are in blue while outgoing links are in red.

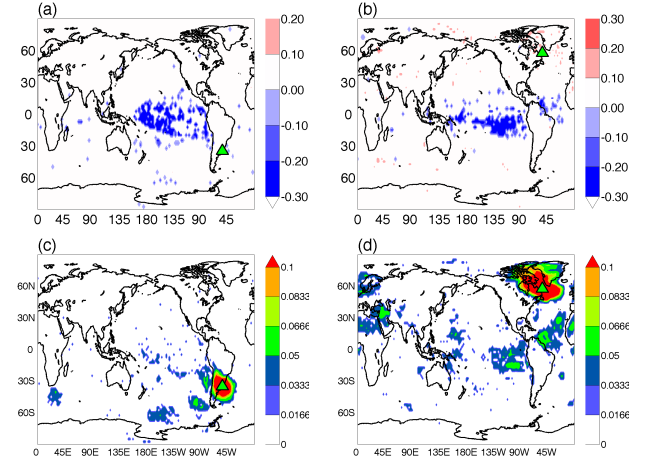


FIG. 4. (Color Online). (a), (b) Directionality of the significant links ($\tau = 1$ month) of two nodes in the extratropics indicated with triangles: (a) in southern South America (de la Plata basin) and (b) in Labrador Sea. As the decorrelation time on the extratropics is very fast, no clear structures are seen. In order to show that it is not a problem of statistics, panels (c) and (d) display the significant MI values.

tions decreases, suggesting that the time-scale of decorrelation of the SAT is less than 6 months. This is consistent with the persistence time scale of 3 to 6 months of observed sea surface temperature.

The extratropical atmosphere shows larger internal variability than the tropics and the impact of the extratropical SST on the atmosphere is much more limited than in the tropics. Thus, the variability of extratropi-

cal SAT is dominated by synoptic atmospheric dynamics and has time scales of a few days. Longer persistence time scales might appear in the extratropics if the region is influenced by tropical SST. This motivates the use of a small value of τ when considering extra-tropical variability.

In Fig. 4 the DI and MI maps for two points in the extratropics are shown. Panels 4 (a,c) show links related to a point in southeastern South America, while panels 4 (b,d) show links related to a point in the Labrador sea, whose characteristics are linked to the North Atlantic Oscillation (NAO)[35]. The top panels show DI for $\tau = 1$ month while the bottom panels show MI.

Consistent with the previous description, the extra-tropical SAT show only some incoming links from the tropical region for $\tau = 1$ month. The point over the Labrador Sea seems to show also some outgoing links to the northeast, although there is no clear structure. As shown in Sec. III C, a better identification of link directionality is obtained by using daily data.

2. Comparison between monthly and daily datasets

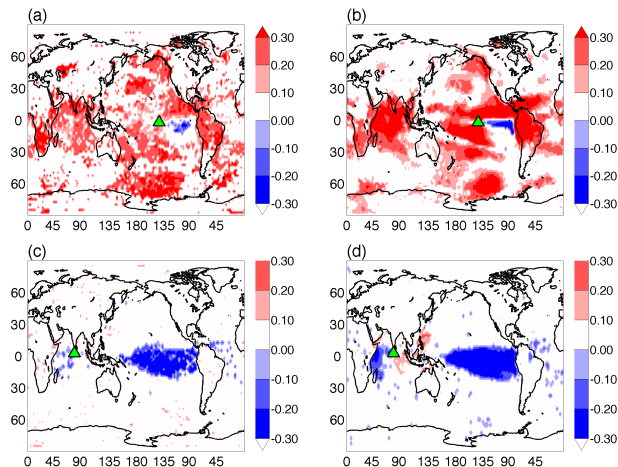


FIG. 5. (Color Online). Comparison of monthly-averaged and daily-averaged datasets. The nodes considered are as in Fig. 1. Panels (a) and (c) display the DI measure ($\tau = 1$ month) calculated from monthly data—same as Figs. 1 (c) and (f). Panels (b) and (d) display the DI measure ($\tau = 30$ days) calculated from daily data. Results are consistent and the resolution using daily data is better.

In order to obtain more temporal resolution daily data has been used. Figure 5 shows a comparison between DI for monthly data—panels 5 (a) and 5 (c)—and for daily data—panels 5 (b) and 5 (d). Corresponding to the same points considered in Fig. 1. In order to adequately compare the datasets, τ was chosen equal to one month in the monthly data and 30 days in the daily data.

The maps using monthly and daily data show similar

features and no inconsistencies are found. Furthermore, the map constructed using daily data captures much better the local and remote dependencies and directionality of the links. Areas with significant links are better defined and some regions that are known to be influenced by equatorial Pacific SST, like the tropical north Atlantic [36], clearly appear using daily data, but only very roughly using monthly data. Thus, the increase in temporal resolution improves the representation of the links related to tropical regions.

C. Analysis of daily-averaged SAT anomalies

1. Influence of the parameter τ

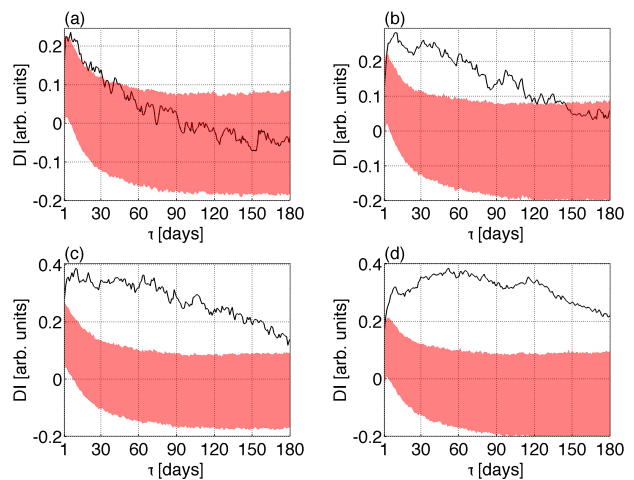


FIG. 6. (Color Online). Influence of the parameter τ when DI is computed from daily-averaged data. The DI value is plotted *vs.* τ for the four links that connect a node in the Pacific ocean with (a) a node in Labrador sea, (b) a node in tropical north Atlantic (in the red area in Fig. 5 between the Caribbean and Africa), (c) a node in southern Pacific (the red spot in Fig. 5 near Antarctica) and (d) a node in the Indian ocean. In order to filter the noise the DI value was averaged over second neighbors in the grid. The shaded area (red online) indicates the DI value computed from 100 BS surrogates, as explained in section III A.

To analyze the influence of the parameter τ in the DI measure when it is computed from daily data, four links connecting the node in the Pacific ocean are considered in Figs. 1, 5 with:

- (a) a node in Labrador sea,
- (b) a node in tropical north Atlantic (in the red area in Fig. 5 between the Caribbean and Africa),
- (c) a node in southern Pacific (the red spot in Fig. 5 near Antarctica) and
- (d) a node in the Indian ocean.

Figure 6 displays, for each link, the DI value *vs.* τ .

In panel (a), the directionality index that characterizes the influence of the equatorial Pacific to the Labrador

Sea shows a small increase for small τ that persists up to about $\tau = 10$ days. This is the typical time scale associated with the setup of the atmospheric anomalies forced by anomalies in tropical convection. Afterwards, the DI decreases exponentially-like, becoming non-significant at $\tau = 30$ days—as the DI value enters the shaded area. In the tropical north Atlantic—Fig. 6 (b)—the DI to the equatorial Pacific shows similar values as for the Labrador Sea case, for up to $\tau = 10$ days. For larger values of τ , however, the DI has significant values (approximately constant) up to about $\tau = 60$ days. Afterwards, it decreases becoming non-significant for τ larger than 4 months. The difference in behavior between the tropical north Atlantic and the Labrador Sea might be because the remote forcing from ENSO induces a clear regional response in the surface temperatures of the tropical Atlantic [36], which will add persistence to the remote signal. On the other hand, in the Labrador Sea, the ocean does not respond and the large atmospheric variability obscures the signal from the equatorial Pacific.

A similar behavior to the one seen in the tropical Atlantic is also found in the Southern ocean—Fig. 6 (c). There is a fast time scale for small values of τ , but in this case the influence of the equatorial Pacific persists for τ up to 80 days, with DI values significantly larger than in the Atlantic. The ENSO influence over the south Pacific is one of the most robust signals in the extratropics, consequence of atmospheric teleconnections associated with the Pacific South American pattern [37]. The time scale of about 3 months seen in the DI is likely associated with the time it takes to the surface ocean to respond to anomalous atmospheric fluxes and to the seasonal dependence of the atmospheric teleconnection pattern on the mean state of the extratropical atmosphere.

The behavior of the influence of the equatorial Pacific onto the Indian ocean—Fig. 6 (d)—also shows a fast time scale of a few days, but in this case the largest value of the DI is seen for a τ of about 60 days. Also, there are large values of DI for τ values of more than 4 months. This suggests that the Indian ocean responds to the incoming ENSO signal in a time scale of about 2 months, through thermodynamic and dynamic coupling [38, 39].

Thus, the DI plots as a function of τ for different ocean regions suggest that there is one fast time scale of about 10 days associated to the setup of the atmospheric teleconnections associated with changes in tropical convection. Moreover, in some regions the DI shows a second longer time scale of about 2 or 3 months associated with the response of the local ocean to the circulation anomalies forced from the tropics. Finally, longer time scales related to oceanic inertia also affect the DI value but longer datasets are needed in order to perform a robust estimation of their effects.

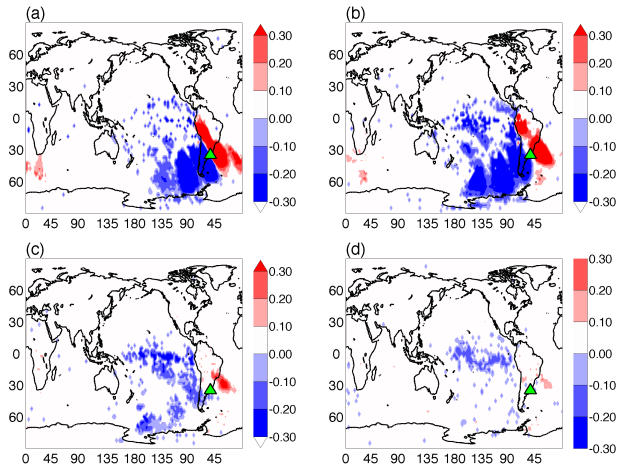


FIG. 7. (Color Online). Effect of τ in the southern extratropics, when DI is computed from daily data. The same node in Fig. 4 (a) is considered, and the values of τ are: (a) 1 day, (b) 3 days, (c) 7 days, and (d) 30 days. Small τ capture wave trains propagating from west to east, while for τ values longer than a week the decorrelation is higher and only the influence from the Pacific ocean persists.

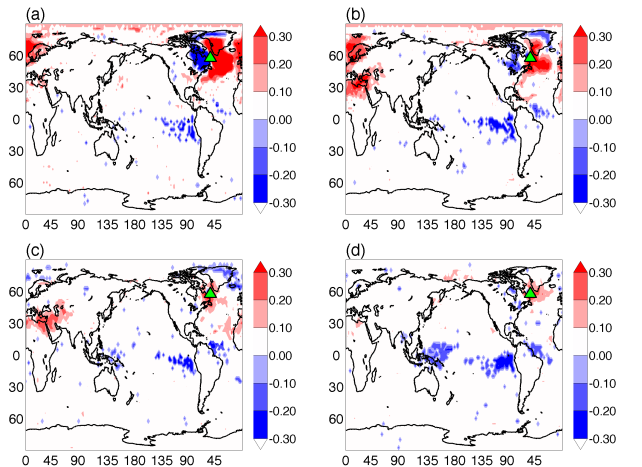


FIG. 8. (Color Online). As in Fig 7 but for a node in the northern extratropics —the Labrador Sea, as in Fig. 4 (b). The values of τ are as in Fig 7. The Labrador sea area is related to a source of atmospheric variability of the north Atlantic ocean that affects Europe. As in Fig. 7, smaller values of τ can capture wave trains propagating from west to east, over Europe, while values of τ longer than a week loose this effect and only the influence from the Pacific ocean persists.

2. Influence of τ in the extra-tropics

The improvement in characterizing directionality using daily data is even larger in the extratropics. As mentioned above, the extratropical SAT is strongly dependent on synoptic scale perturbations (a time scale of a

few days). Thus, the use of daily data should allow to uncover these relationships and investigate the direction of the links as the lag increases. In order to do so, the point in southeastern South America shown in Fig. 4 is considered and the directionality network for several values of τ ranging from a day to one month (Fig. 7) is constructed. For synoptic time scales of a few days the methodology uncovers the existence of a wave train connected to southeastern South America propagating with a southwest-northeast direction. Moreover, there is a clear separation line between regions with incoming and outgoing links. This configuration is characteristic of the progression of a front through the reference point and does not imply that the SAT over the reference point influences the region to the northeast but it only happens to be in the path of the perturbation. As the lag time increases, the extratropical wave train associated with synoptic time scales fades and only the points in the tropics remain, consistent with an influence of the equatorial Pacific on the region on longer time scales, perhaps related to ENSO.

A similar behavior is seen taking as reference point the SAT over the Labrador Sea (Fig. 8). For small values of τ the progression of a front is clearly detected using this procedure: given the mean westerly winds at these latitudes, the front moves from west to east and is clearly marked as the boundary between the incoming and outgoing links. It is also seen for $\tau = 3$ suggesting that in about three days the front reaches the Mediterranean region affecting temperatures there. Again, as τ increases mainly the tropical links remain. However, even for $\tau = 30$ there is a well defined region of outgoing links that remain over the Labrador Sea, suggesting that the SAT in the region may have relatively long time scales of variability, perhaps related to the North Atlantic Oscillation. Another possibility is the influence of the local ocean that increases the persistence of atmospheric temperature anomalies through thermodynamic coupling, as shown by Barsugli *et al.* [40].

3. Influence of τ in the tropics

Figure 9 presents the effect of τ on daily data in the tropics. Panel 9 (a) shows the point considered in the Pacific ocean for $\tau = 1$ day. Contrarily to e.g. figure 7, where front propagation shows a big response in the tropics due to the existence of wave trains in the extratropics, in this case the connections are weak. The opposite case, for $\tau > 30$ days—Panels 9 (b,c)—are consistent with in figure 6 as the structures are robust for a wide range of τ . Notice in Panel 9 (c) the presence of a second blue area near the coast of Chile, This is in agreement with the easterly trades advection over the Pacific ocean. In order to complement the information provided in panels 9 (a,b,c) and also in Fig. 6, in the supplementary information a movie was included [41] (Multimedia view), which presents the links to the node in the Pacific ocean,

for $\tau = 1$ day to $\tau = 180$ days. Panels 9 (d,e,f) show maps for a point in the Indian ocean, for values of $\tau = 1, 45$, and 90 days respectively where similar results are obtained.

4. Directionality on the tropical Pacific Ocean

In Fig. 10 the DI for $\tau = 30$ days has been plotted for points covering all the equatorial Pacific. They begin at 95.0° W—panel 10 (a)—and end at 125.0° E—panel 10 (d). Clearly, the influence the Pacific ocean exerts is almost global, over tropics and extratropics, in agreement with previous studies (e.g. in [42]). The DI allows to show that—even if there are feedbacks and the Pacific is affected by extratropical perturbations and other ocean basins (e.g. the tropical Atlantic)—the influence is *effectively* from the Pacific to the rest of the world. Moreover, the maps show that the largest influence is done by the equatorial Pacific close to the dateline. This is clear in the extratropical atmosphere, as well as in the tropical north Atlantic. On the other hand, the connection to the Indian ocean and south Atlantic is not so sensitive of the point considered over the equatorial Pacific. As the reference point moves further west from the dateline the influence decreases substantially, only remaining weak connections to the tropical north Atlantic and Indian oceans. The methodology can thus be applied to find the best region to construct an index that describes the Pacific influence over the area where climate anomalies are studied.

Notice that all maps show a blue tongue of incoming links to the east of the point considered; it is seen first in panel 10 (a) and extends westward until covering the whole Pacific ocean in 10 (d). This feature is related to the existence of the equatorial cold tongue and the fact that easterly trades blow over the equator thus advecting air from the east to the west of the point.

IV. CONCLUSIONS

The climatological relevance of *directed* climate networks constructed via information-theoretic tools has been shown. The presence of significant links was inferred using mutual information while the direction of those links was inferred using the directionality index. Using monthly-averaged SAT data the results from our previous work [33] have been recovered and the analysis extended to infer the net direction of propagation of the information. The inference method was tested against the value of the parameter τ , that represents the time required for the information to travel from one node to another. It was found that by adequately tuning τ the network connectivity varied revealing the various time-scales of atmospheric processes; too short values of τ failed to capture several long-range links, while for too large values of τ (above the length of correlations in the

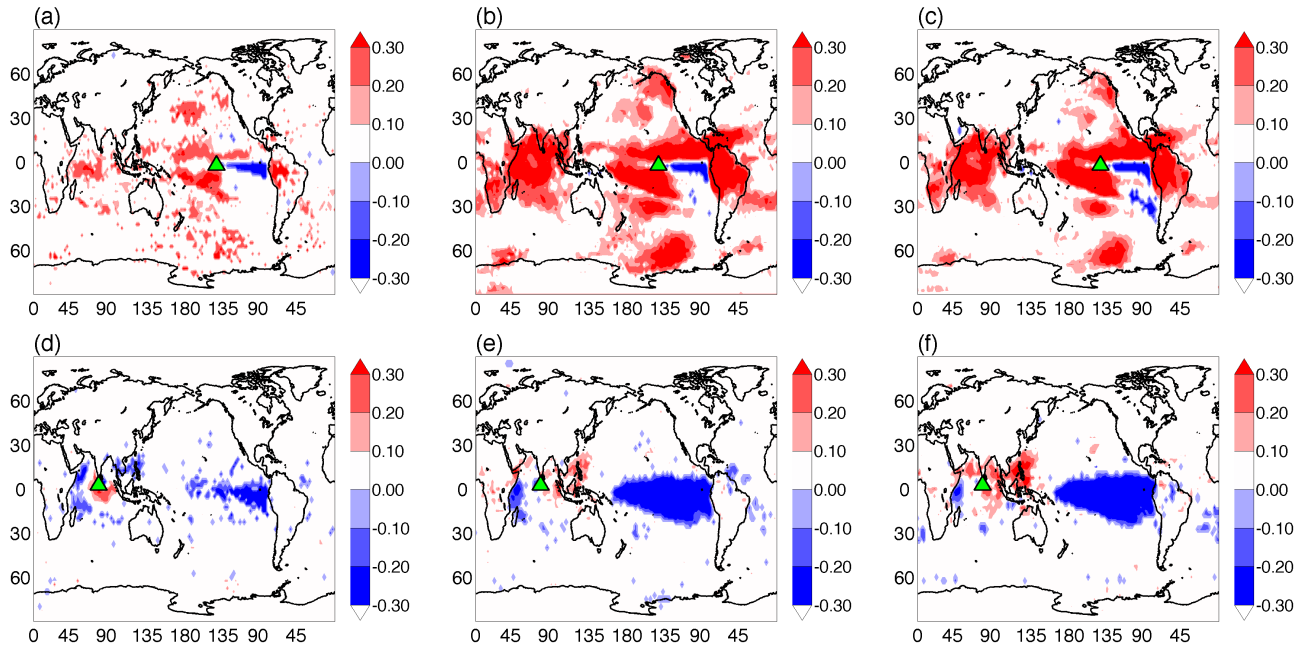


FIG. 9. (Multimedia view) (Color Online). Effect of τ using daily data for a node in the tropics. In panels (a-c) the node is in the Pacific ocean, and (d-f), in the Indian ocean. The values of τ are: (a,d) 1 day, (b,e) 45 days, and (c,f) 90 days. The maps for $\tau = 30$ were shown in Fig. 3. In contrast to the extra-tropics, where the propagation of waves in the scale of days is dominant, and are clearly seen in the maps for time scales of a few days, in the tropics the variability is longer as the ocean adds persistence to the nearly barotropic atmosphere. For higher values of τ only long lasting phenomena are observed, related to the strongest variability pattern in this area, which is ENSO. In the supplementary information a movie showing the links to the node of the first row, for $\tau = 1$ to $\tau = 180$ is presented [41].

data) the connectivity of the network decreased drastically.

In addition, when considering daily-averaged SAT data, the analysis revealed variability patterns consistent with known features of the global climate dynamics. In the extratropics the long time average synoptic weather was correctly inferred: as specific examples two geographical regions in different hemispheres—one node in de la Plata basin and another in the Labrador Sea—were considered, and the link direction revealing wave trains propagating from west to east, in both hemispheres was shown.

As a future work, it would be interesting to apply this methodology to analyze how the climate network changes

in the different seasons. Another interesting issue is the role of permutation entropy. Computing the directionality index using the ordinal patterns approach [32, 33], which would allow constructing networks revealing atmospheric processes with short time-scales (of days-months) or with long time-scales (few years). These studies are in progress and will be reported elsewhere.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2011 under grant agreement no 289447 (ITN LINC). C.M. also acknowledges financial support of grant FIS2012-37655-C02-01 of the Spanish Ministerio de Ciencia e Innovación.

-
- [1] D. J. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
 - [2] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
 - [3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, *Phys. Rep.* **424**, 175 (2006).
 - [4] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Phys. Rep.* **469**, 93 (2008).
 - [5] V. Colizza, A. Barrat, M. Barthélemy, and A. Vespignani, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2015 (2006).
 - [6] A. L. Barabási, H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek, *Phys. A Stat. Mech. its Appl.* **311**, 590 (2002).
 - [7] R. Albert, H. Jeong, and A. L. Barabási, *Nature* **401**, 398 (1999).

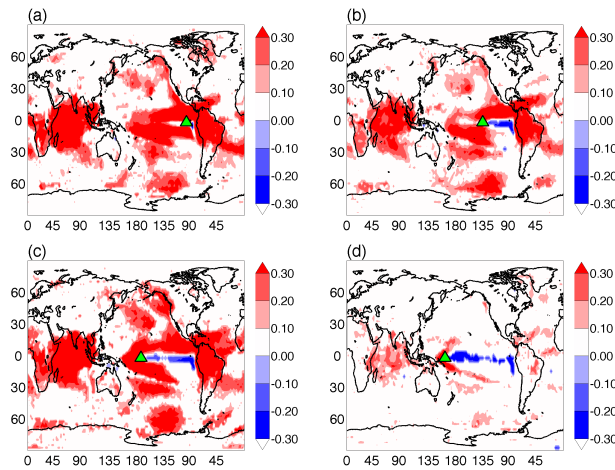


FIG. 10. (Color Online). The zonal change of directionality over the equatorial Pacific is shown. In all cases $\tau = 30$ days. From (a) near the south American coast, to (d) in the western Pacific ocean. As seen in the maps, most of the points over central and eastern Pacific have an important effect over a large part of the world; especially over the tropical areas and the rest of the Pacific ocean—notice that in (a)-(c) the teleconnections remain basically the same, although the intensity varies. Incoming links are in blue while outgoing are in red.

- [8] A. A. Tsonis and P. Roebber, Phys. A Stat. Mech. its Appl. **333**, 497 (2004).
- [9] K. Yamasaki, K. J. Mackin, M. Ohshiro, K. Matsushita, and E. Nunohiro, Artif. Life Robot. **12**, 122 (2008).
- [10] K. Steinhaeuser, N. V. Chawla, and A. R. Ganguly, ACM SIGKDD Explor. Newsl. **12**, 25 (2009).
- [11] T. Zerenner, P. Friederichs, K. Lehnertz, and A. Hense, Chaos **24**, 023103 (2014).
- [12] S. Bialonski, M.-T. Horstmann, and K. Lehnertz, Chaos **20**, 013134 (2010), arXiv:arXiv:1105.2257v1.
- [13] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Eur. Phys. J. Spec. Top. **174**, 157 (2009).
- [14] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, EPL **87**, 48007 (2009).
- [15] J. F. Donges, H. C. H. Schultz, N. Marwan, Y. Zou, and J. Kurths, Eur. Phys. J. B **84**, 635 (2011).
- [16] A. A. Tsonis, K. L. Swanson, and G. Wang, J. Clim. **21**, 2990 (2008).
- [17] A. Tantet and H. A. Dijkstra, Earth Syst. Dyn. **5**, 1 (2014).
- [18] M. van der Mheen, H. A. Dijkstra, A. Gozolchiani, M. den Toom, Q. Feng, J. Kurths, and E. Hernandez-Garcia, Geophys. Res. Lett. **40**, 2714 (2013).
- [19] A. Gozolchiani, K. Yamasaki, O. Gazit, and S. Havlin, EPL **83**, 28005 (2008).
- [20] K. Yamasaki, A. Gozolchiani, and S. Havlin, Phys. Rev. Lett. **100**, 228501 (2008).
- [21] G. Wang, P. Yang, X. Zhou, K. L. Swanson, and A. A. Tsonis, Geophys. Res. Lett. **39**, 704 (2012).
- [22] I. I. Mokhov, D. a. Smirnov, P. I. Nakonechny, S. S. Kozlenko, E. P. Seleznev, and J. Kurths, Geophys. Res. Lett. **38**, L00F04 (2011).
- [23] K. Hlavackovaschindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, Phys. Rep. **441**, 1 (2007).
- [24] M. Paluš, Phys. Rev. Lett. **112**, 078702 (2014).
- [25] J. Hlinka, D. Hartman, M. Vejmelka, J. Runge, N. Marwan, J. Kurths, and M. Paluš, Entropy **15**, 2023 (2013).
- [26] A. Bahraminasab, F. Ghasemi, A. Stefanovska, P. McClintock, and H. Kantz, Phys. Rev. Lett. **100**, 084101 (2008).
- [27] M. Paluš and A. Stefanovska, Phys. Rev. E **67**, 055201 (2003).
- [28] K. Lehnertz and H. Dickten, Proc. R. Soc. A **in press** (2014).
- [29] E. Kalnay *et al.*, Bull. Am. Meteorol. Soc. **77**, 437 (1996).
- [30] M. Paluš, Contemp. Phys. **48**, 307 (2007).
- [31] J. Amigó, *Permutation Complexity in Dynamical Systems*, Springer Series in Synergetics (Springer Berlin Heidelberg, Berlin, Heidelberg, 2010) p. 260.
- [32] M. Barreiro, A. C. Marti, and C. Masoller, Chaos **21**, 013101 (2011).
- [33] J. I. Deza, M. Barreiro, and C. Masoller, Eur. Phys. J. Spec. Top. **222**, 511 (2013).
- [34] M. Mudelsee, *Climate Time Series Analysis*, Atmospheric and Oceanographic Sciences Library, Vol. 42 (Springer Netherlands, Dordrecht, 2014).
- [35] J. I. Deza, C. Masoller, and M. Barreiro, Nonlinear Process. Geophys. **21**, 617 (2014).
- [36] P. Chang, R. Saravanan, L. Ji, and G. Hegerl, J. Clim. **13**, 2195 (2000).
- [37] K. C. Mo and R. W. Higgins, Mon. Weather Rev. **126**, 1581 (1998).
- [38] H. Annamalai, R. Murtugudde, J. Potemra, S. Xie, P. Liu, and B. Wang, Deep Sea Res. Part II Top. Stud. Oceanogr. **50**, 2305 (2003).
- [39] X. Wang and C. Wang, Clim. Dyn. **42**, 991 (2013).
- [40] J. Barsugli and D. Battisti, J. Atmos. Sci. **55**, 477 (1998).
- [41] “See supplemental material at <http://youtu.be/alcormjKIbM> for a movie showing an DI maps for increasing τ . year = 2013.”
- [42] K. E. Trenberth, G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewsky, J. Geophys. Res. **103**, 14291 (1998).